63-1-4

D1-82-0197

BOEING SCIENTIFIC RESEARCH LABORATORIES

Elementary Relations Between Uniform and Normal Distributions in the Plane

George Marsaglia

Mathematics Research

August 1962

D1-82-0197

ELEMENTARY RELATIONS BETWEEN UNIFORM AND NORMAL DISTRIBUTIONS IN THE PLANE

by

George Marsaglia

Mathematical Note No. 268

Mathematics Research Laboratory

BOEING SCIENTIFIC RESEARCH LABORATORIES

August 1962

0. Summary

If (U,V) is uniformly distributed over the unit circle, then $U^2 + V^2$ is uniformly distributed over $[0,1]$ and is independent of U/V. This simple result leads to an improved method for generating normal random variables via polar coordinates, and also serves to relate normal probability measure and Lebesque measure.

1. Main Paragraph

Let (U,V) be uniformly distributed over the unit circle: $\{(u,v): u^2 + v^2 \leq 1\}$. <u>Then</u> $U^2 + V^2$ <u>is uniformly distributed over</u> $[0,1]$ <u>and is independent of</u> U/V. The proof is simple and omitted. It follows that $U^2 + V^2$ is independent of $[1 + (V/U)^2]^{-\frac{1}{2}}$ and of $[1 + (U/V)^2]^{-\frac{1}{2}} = V(U^2 + V^2)^{-\frac{1}{2}}$, and hence

$$(1)\qquad \begin{aligned} X &= U\left[\frac{-2\ln(U^2 + V^2)}{U^2 + V^2}\right]^{\frac{1}{2}} \\ Y &= V\left[\frac{-2\ln(U^2 + V^2)}{U^2 + V^2}\right]^{\frac{1}{2}} \end{aligned}$$

are independent standard normal random variables, since $U(U^2 + V^2)^{-\frac{1}{2}}$ and $V(U^2 + V^2)^{-\frac{1}{2}}$ are distributed as $\cos\theta$ and $\sin\theta$, and $[-2\ln(U^2 + V^2)]^{\frac{1}{2}}$ is distributed as ρ, in the polar representation (ρ,θ) of (X,Y).

2. Generating Normal Variables

In [1], Box and Muller point out the elementary fact that a pair of normal random variables may be produced in a computer by generating ρ and Θ, then putting $X = \rho \cos \Theta$ and $Y = \rho \sin \Theta$. The obvious method - use two independent uniform $[0,1]$ random variables (U_1, U_2) and put $X = \cos(2\pi U_1)(-2 \ln U_2)^{\frac{1}{2}}$, $Y = \sin(2\pi U_1)(-2 \ln U_2)^{\frac{1}{2}}$, requires cosine, sine, logarithm, and square root subroutines. This method is not very practical unless some tricks are used to speed it up. One line of improvement is to generate $\cos \Theta$ and $\sin \Theta$ in the form $U(U^2 + V^2)^{-\frac{1}{2}}$ and $V(U^2 + V^2)^{-\frac{1}{2}}$, as suggested by von Neumann [2]. The representation in (1) goes one step further, taking advantage of the fact that $U^2 + V^2$ is independent of $U(U^2 + V^2)^{-\frac{1}{2}}$ and of $V(U^2 + V^2)^{-\frac{1}{2}}$, is itself uniform $[0,1]$, and may be used to form ρ by way of $\rho = [-2 \ln(U^2 + V^2)]^{\frac{1}{2}}$.

Although there are faster methods for generating normal variables, see, e.g., [3] - [5], the method suggested by relations (1) may be suitable for situations where ease of programming is the primary consideration. Furthermore, a slight modification of the procedure may be used to dispose of the problem of handling the tail of the normal distribution in one of the super-fast programs. Let $r > 0$ be a constant. Then putting

$$(2)\qquad \begin{aligned} X &= U\left[\frac{r^2 - 2\ln(U^2 + V^2)}{U^2 + V^2}\right]^{\frac{1}{2}} \\ Y &= V\left[\frac{r - 2\ln(U^2 + V^2)}{U^2 + V^2}\right]^{\frac{1}{2}} \end{aligned}$$

will produce a normal pair (X,Y) conditioned by $X^2 + Y^2 \geq r^2$. If we want to produce a normal variable Z conditioned by $|Z| \geq r$, we may generate X,Y according to (2) and put $Z = X$ if $|X| \geq r$. If $|X| < r$, test: is $|Y| \geq r$? If yes, put $Z = Y$, if no, generate a new pair X,Y and try again.

3. Normal Measure → Lebesque Measure

Finally, expression (1) is convenient for relating normal probability measure and Lebesque measure in the plane. Let T^{-1} be the mapping of the unit circle $C: \{(u,v): 0 < u^2 + v^2 \leq 1\}$ onto the plane, R_2, of points (x,y) given by the relations in (1):

$$x = u\left[\frac{-2\ln(u^2+v^2)}{u^2+v^2}\right]^{\frac{1}{2}}$$

$$y = v\left[\frac{-2\ln(u^2+v^2)}{}\right]^{\frac{1}{2}} .$$

Then T^{-1} is one-to-one, and its inverse, T, maps points (x,y) of R_2 into points (u,v) of C according to the relations:

$$u = x\left[\frac{e^{-\frac{1}{2}(x^2+y^2)}}{x^2+y^2}\right]^{\frac{1}{2}}$$

(3)

$$v = y\left[\frac{e^{-\frac{1}{2}(x^2+y^2)}}{x^2+y^2}\right]^{\frac{1}{2}} .$$

Let μ be standard normal probability measure in the plane, and λ Lebesque measure. Then

$$\mu(A) = P[(X,Y) \in A] = P[(U,V) \in T(A)] = \frac{\lambda(T(A))}{\pi} .$$

That is, we may find the standard normal probability measure of regions A by mapping A into the unit circle according to relations (2), then finding the area of $T(A)$.

This type of transformation has been used before, [6]. We have a simpler version here, which may be put in this form: To find the standard normal probability measure of a region A, map each polar coordinate point (ρ,θ) of A into the polar coordinate point $(e^{-\frac{1}{4}\rho^2},\theta)$ of the unit circle, and find $\frac{1}{\pi}$ times the area of the transformed region. Inverting the mapping $X = \cos 2\pi U_1(-2 \ln U_2)^{\frac{1}{2}}$, $Y = \sin 2\pi U_1(-2 \ln U_2)^{\frac{1}{2}}$ gives a more suitable method for converting normal to Lebesque measure, however. The procedure runs as follows: To find the normal probability measure of a region A, map each polar coordinate point (ρ,θ) of A into the point (u,v) of the unit square $\{(u,v): 0 < u < 1, 0 < v < 1\}$ according to the relations

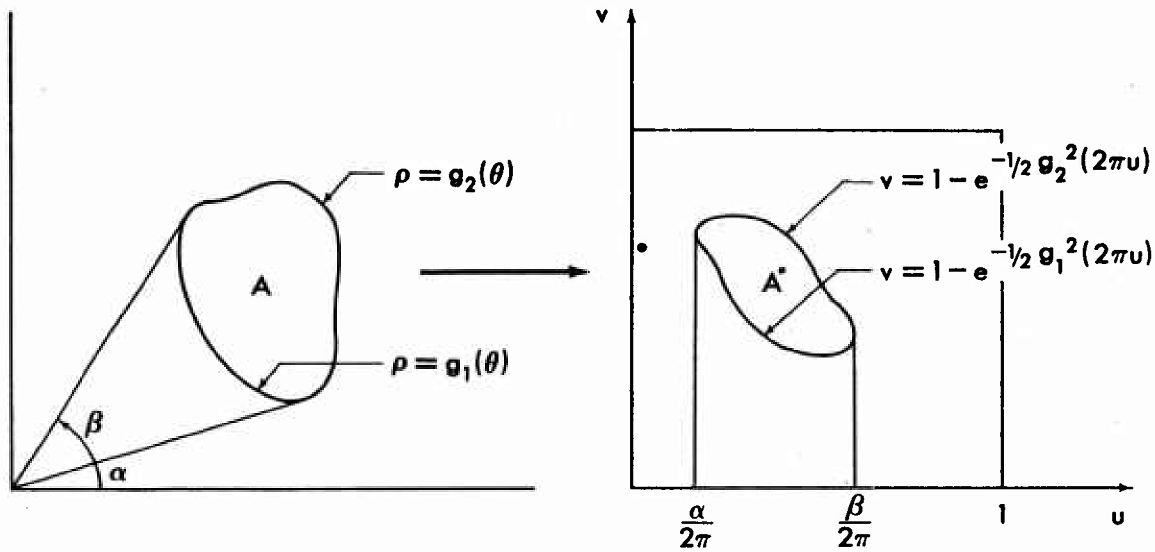
$$u = \theta/2\pi$$

(4)

$$v = 1 - e^{-\frac{1}{2}\rho^2}$$

then find the area of the transformed region. We have found relations (4) to be the most suitable for a general purpose procedure for transforming normal to Lebesque measure, the advantage being that it is relatively easy to assign an equally spaced set of points $u_0, u_1, \ldots$ and hence use one of the standard numerical integration procedures.

The transformation may be described graphically in this way:



Use of Simpson's rule gives the approximation

$$\mu(A) = \lambda(A^*) \cong \frac{\beta - \alpha}{12\pi n}[t_0 + 4t_1 + 2t_2 + 4t_3 + \cdots + 2t_{2n-2} + 4t_{2n-1} + t_{2n}]$$

where $t_i = e^{-\frac{1}{2}g_1^2(2\pi u_i)} - e^{-\frac{1}{2}g_2^2(2\pi u_i)}$, $u_i = \alpha/2\pi + i\sigma$, and $\sigma = (\beta - \alpha)/4\pi n$.

This elementary procedure makes it possible to write a single computer program which will handle a large number of the commonly encountered regions in the normal probability plane - polygonal regions, ellipses, intersections of ellipses, etc.

REFERENCES

[1] G. E. P. Box and M. E. Muller, A note on the generation of normal deviates. Ann. Math. Stat. 38, (1958), pp. 610-611.

[2] J. von Neumann, Various techniques used in connection with random digits. Monte Carlo Method. National Bureau of Standards AMS 12 (1951), pp. 167-174.

[3] G. Marsaglia, Random Variables and Computers, Transactions of the Third Prague Conference, Prague, 1962. (to appear). Also available as Boeing Scientific Research Laboratories Document D1-82-0182.

[4] G. Marsaglia, Expressing a random variable in terms of uniform random variables, Ann. Math. Stat.32 (1961), pp. 894-898.

[5] G. Marsaglia, M. D. MacLaren, T. A. Bray, Procedures for Generating Normal Random Variables II, Boeing Scientific Research Laboratories Document No. ____, August 1962.

[6] G. R. Van Brocklin, Jr. and P. G. Murray, A polar-planimeter method for determining the probability of hitting a target, Operations Research 4; (1956), pp. 87-91.